ASYMPTOTIC PROPERTIES OF MAXIMUM LIKELIHOOD ESTIMATORS BASED ON CONDITIONAL SPECIFICATION

by

P. K. Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1158

February 1978

ASYMPTOTIC PROPERTIES OF MAXIMUM LIKELIHOOD ESTIMATORS BASED ON CONDITIONAL SPECIFICATION*

Pranab Kumar Sen
University of North Carolina, Chapel Hill

ABSTRACT

Along with the asymptotic distribution, expressions for the asymptotic bias and asymptotic dispersion matrix of the preliminary test maximum likelihood estimator for a general multi-sample parametric model (when the null hypothesis relating to the restraints on the parameters may not hold) are derived and compared with the parallel expressions for the unrestricted and restricted maximum likelihood estimators. This study reveals the robustness property of the preliminary test estimator when the assumed restraints may not hold.

AMS 1970 Classification No. 62A10, 62F99

Key Words & Phrases: Asymptotic bias, asymptotic dispersion matrix, asymptotic relative efficiency, conditional specification, maximum likelihood (restricted and unrestricted) estimators, preliminary test estimators, restraints on parameters.

---

*Work supported partially by the National Institutes of Health, Contract No. NIH-HHLB1-71-2243 from the National Heart, Lung and Blood Institute.

1. INTRODUCTION

In a parametric model, assuming that the underlying distributions are of specified forms and the parameter (vector) $\underset{\sim}{\theta}$ belongs to a suitable *parameter space* Ω, the (*unrestricted*) *maximum likelihood estimator* (MLE) $\tilde{\underset{\sim}{\theta}}$ of $\underset{\sim}{\theta}$ is obtained by maximizing (over $\underset{\sim}{\theta} \in \Omega$) the *likelihood function* of the sample observations. Under appropriate regularity conditions, $\tilde{\underset{\sim}{\theta}}$ is an (asymptotically) optimal estimator of $\underset{\sim}{\theta}$. In certain problems, ω, a proper sulspace of Ω, can be identified from extraneous considerations and a *restricted MLE* $\hat{\underset{\sim}{\theta}}$ of $\underset{\sim}{\theta}$ can be derived by maximizing the likelihood function subject to the restrain that $\underset{\sim}{\theta} \in \omega$. When $\underset{\sim}{\theta} \in \omega$, $\hat{\underset{\sim}{\theta}}$ is (asymptotically) a better estimator then $\tilde{\underset{\sim}{\theta}}$. But, if contrary to this assumed restrain, actually $\underset{\sim}{\theta} \notin \omega$, then $\hat{\underset{\sim}{\theta}}$ may not only loose its optimality but also may be a biased (or even an inconsistent) estimator. This lack of validity-robustness of $\hat{\underset{\sim}{\theta}}$ may be of some concern in a class of problems arising in applied statistics, where $\omega(\subset \Omega)$ can be suggested from certain practical considerations, but, there may not be sufficient apriori evidence of $\underset{\sim}{\theta} \in \omega$ so as to warrant the use of $\hat{\underset{\sim}{\theta}}$ without any reservation. In such a case, a compromise between $\hat{\underset{\sim}{\theta}}$ and $\tilde{\underset{\sim}{\theta}}$ based on a *conditional specification* appears to be appealing: A *preliminary* (*likelihood-ratio*) *test* for H_0: $\underset{\sim}{\theta} \in \omega$ is made, the preliminary test (PT) estimator $\underset{\sim}{\theta}^*$ is then taken to be $\hat{\underset{\sim}{\theta}}$ or $\tilde{\underset{\sim}{\theta}}$ according as H_0: $\underset{\sim}{\theta} \in \omega$ is tenable or not.

For a variety of specific problems, mostly, relating to univariate and multivariate normal distributions, various workers have considered various PT estimators; we may refer to Kitagawa (1963) and a recent bibliography by Bancroft and Han (1977). The object of the present investigation is to study the asymptotic properties of the PTMLE $\underset{\sim}{\theta}^*$ under the classical regularity conditions pertaining to the asymptotic theory of $\hat{\underset{\sim}{\theta}}$ or $\tilde{\underset{\sim}{\theta}}$ [viz., Aitchison and Silvey (1958)]. As is usually the case with PT estimators, $\underset{\sim}{\theta}^*$ is not generally (asymptotically) optimal or unbiased when $\underset{\sim}{\theta} \in \omega$; nevertheless, it has good asymptotic properties when $\underset{\sim}{\theta} \in \omega$. For $\underset{\sim}{\theta} \notin \omega$, $\underset{\sim}{\theta}^*$ may perform better than either of $\hat{\underset{\sim}{\theta}}$ and $\tilde{\underset{\sim}{\theta}}$. Indeed, for $\underset{\sim}{\theta}$ close to the boundary of ω, expressions for the asymptotic bias and dispersion matrix have been derived for each of $\hat{\underset{\sim}{\theta}}$, $\tilde{\underset{\sim}{\theta}}$ and $\underset{\sim}{\theta}^*$ and incorporated in the study of the comparative performances of these estimators.

Along with the preliminary notions, these estimators are introduced in Section 2. Section 3 deals with their asymptotic distributions when $\underset{\sim}{\theta} \in \omega$. Parallel results for the non-null case are presented in Section 4. The last section is devoted to the asymptotic comparisons of these estimators.

2. PRELIMINARY NOTIONS

Since in a PTE problem, typically, a multi-sample situation may be invovled, we conceive of $k(\geq 1)$ independent samples and let $X_{i1},\ldots,X_{in_i}$ by n_i independent and identically distributed random vectors (i.i.d.r.v.) with a distribution function (df) $F_i(x,\underset{\sim}{\theta})$, for $i = 1,\ldots,k$, where $x \in E^p$, the $p(\geq 1)$-dimensional Euclidean space and $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_t)' \in \Omega \subset E^t$. Actually, F_i may not depend on all the parameters $\theta_1,\ldots,\theta_t$ for every $i(= 1,\ldots,k)$, rather, each element of $\underset{\sim}{\theta}$ is associated with at least one df. Further, we assume that for each $\underset{\sim}{\theta} \in \Omega$ and $i(= 1,\ldots,k)$, $F_i(x,\underset{\sim}{\theta})$ admits a density function $f_i(x,\underset{\sim}{\theta})$ (with respect to some sigma-finite measure μ). Then, the (log-) likelihood function is defined by

$$(2.1) \qquad \log L_n(\underset{\sim}{X}_n,\underset{\sim}{\theta}) = \sum_{i=1}^{k}\sum_{j=1}^{n_i} \log f_i(X_{ij},\underset{\sim}{\theta}) \;, \quad \underset{\sim}{\theta}\in\Omega \;,$$

where $n = \sum_{n-1}^{k} n_i$ and $\underset{\sim}{X}_n = (X_{11},\ldots,X_{kn_k})$ is the sample-point $(\in E^{pn})$. The true parameter $\underset{\sim}{\theta}_0(\in \Omega)$ is not known. A unrestricted MLE $\underset{\sim}{\tilde\theta}_n$ is an element of Ω such that

$$(2.2) \qquad \log L_n(\underset{\sim}{X}_n,\underset{\sim}{\tilde\theta}_n) = \sup_{\underset{\sim}{\theta}\in\Omega} \log L_n(\underset{\sim}{X}_n,\underset{\sim}{\theta}) \;.$$

Suppose now that $\underset{\sim}{\theta}_0$, though unknown, belongs to a subset ω, where

$$(2.3) \qquad \omega = \{\underset{\sim}{\theta}:\ \underset{\sim}{h}(\underset{\sim}{\theta}) = (h_1(\underset{\sim}{\theta}),\ldots,h_r(\underset{\sim}{\theta})) = \underset{\sim}{0}\} \text{ for some } r < t \;.$$

Then, a restricted MLE $\underset{\sim}{\hat\theta}_n$ is an element of ω such that

$$(2.4) \qquad \log L_n(\underset{\sim}{X}_n,\underset{\sim}{\hat\theta}_n) = \sup_{\underset{\sim}{\theta}\in\omega} \log L_n(\underset{\sim}{X}_n,\underset{\sim}{\theta}) \;.$$

For testing H_0: $\underset{\sim}{\theta}\in\omega$, the classical likelihood ratio statistic is

$$(2.5) \qquad \mathcal{L}_n = -2\log\left\{\left[\sup_{\underset{\sim}{\theta}\in\omega} L_n(\underset{\sim}{X}_n,\underset{\sim}{\theta})\right] \quad \left[\sup_{\underset{\sim}{\theta}\in\Omega} L_n(\underset{\sim}{X}_n,\underset{\sim}{\theta})\right]\right\}$$
$$= -2\log\left(L_n(\underset{\sim}{X}_n;\underset{\sim}{\hat\theta}_n)/L_n(\underset{\sim}{X}_n;\underset{\sim}{\tilde\theta}_n)\right) \;.$$

Let $\ell_{n,\alpha}$ be a real number such that

$$(2.6) \qquad P\{\mathcal{L}_n \geq \ell_{n,\alpha}|H_0\} \geq \alpha > P\{\mathcal{L}_n \geq \ell_{n,\alpha}|H_0\} \;,$$

where $\alpha(0<\alpha<1)$ is the desired level of significance of the test. Then, the likelihood ratio test consists in rejecting H_0: $\theta\in\omega$ when $\mathcal{L}_n > \ell_{n,\alpha}$ and accepting H_0, otherwise. The PTMLE $\underset{\sim}{\theta}^*_n$ is defined by

$$(2.7) \qquad \underset{\sim}{\theta}_n^* = \begin{cases} \hat{\underset{\sim}{\theta}}_n , & \text{if } L_n \le \ell_{n,\alpha} \\ \tilde{\underset{\sim}{\theta}}_n , & \text{if } L_n > \ell_{n,\alpha} . \end{cases}$$

Our primary concern is to study the asymptotic properties of $\{\underset{\sim}{\theta}_n^*\}$ and compare them with those of $\{\hat{\underset{\sim}{\theta}}_n\}$ and $\{\tilde{\underset{\sim}{\theta}}_n\}$, when H_0: $\underset{\sim}{\theta} \in \omega$ may or may not hold.

For our study, we make the following assumptions:

[A1]: Ω is a convex, compact subspace of E^t, and for every $\underset{\sim}{\theta}_1 \ne \underset{\sim}{\theta}_2$, (both $\in \Omega$), for at least one $i(= 1,\ldots,k)$,

$$(2.8) \quad f_i(x,\underset{\sim}{\theta}_1) \ne f_i(x,\underset{\sim}{\theta}_2) , \quad \text{at least on a set of measure non-zero} .$$

[A2]: For every $\underset{\sim}{\theta} \in \Omega$ and every $i(= 1,\ldots,k)$, $Z_i(\underset{\sim}{\theta}) = \int_{E^p} \log f_i(x,\underset{\sim}{\theta}) dF_i(x,\underset{\sim}{\theta}_0)$ exists. In fact, for the ith density, the Kullback-Leibler information is

$$(2.9) \quad I_i(\underset{\sim}{\theta},\underset{\sim}{\theta}_0) = \int_{E^p} \log\left\{f_i(x,\underset{\sim}{\theta}_0)/f_i(x,\underset{\sim}{\theta})\right\} dF_i(x,\underset{\sim}{\theta}_0) = Z_i(\underset{\sim}{\theta}_0) - Z_i(\underset{\sim}{\theta})$$

where for every $\underset{\sim}{\theta} \in \Omega$, $I_i(\underset{\sim}{\theta},\underset{\sim}{\theta}_0) \ge 0$ with the strict equality only when $f_i(x,\underset{\sim}{\theta}) = f_i(x,\underset{\sim}{\theta}_0)$ almost everywhere (a.e.)

[A3]: For every $\underset{\sim}{\theta} \in \Omega$ and $i(= 1,\ldots,k)$, $\log f_i(x,\underset{\sim}{\theta})$ is (a.e.) thrice differentiable with respect to $\underset{\sim}{\theta}$ and

$$(2.10) \quad \left|\left(\partial^s/\partial\theta_a^{s_1}\partial\theta_b^{s_2}\partial\theta_c^{s_3}\right)\log f_i(x,\underset{\sim}{\theta})\right| \le G_s(x) , \quad \forall\ x \in E^p , \quad \underset{\sim}{\theta} \in \Omega$$

where $s_j \ge 0$, $j = 1,2,3$, $s_1 + s_2 + s_3 = s = 1,2,3$ and $1 \le a, b, c \le t$, and where

$$(2.11) \quad \int_{E^p} G_s(x) dF_i(x,\underset{\sim}{\theta}_0) < \infty \quad \text{for } i = 1,\ldots,k \text{ and } s = 1,2,3 .$$

[It is possible to eliminate the third order derivatives conditions in (2.10)-(2.11) by imposing the following:

$$(2.12)\quad \lim_{\delta\downarrow 0}\max_{i,j,\ell}\left\{\left[\sup_{\underset{\sim}{\theta}:\|\underset{\sim}{\theta}-\underset{\sim}{\theta}_0\|<\delta}\left|\frac{\partial^2}{\partial\theta_j\partial\theta_\ell}\log f_i(x,\underset{\sim}{\theta})\right]_{\underset{\sim}{\theta}} - \frac{\partial^2}{\partial\theta_j\partial\theta_\ell}\log f_i(x,\underset{\sim}{\theta})\right]_{\underset{\sim}{\theta}_0}\right|\right]\right\} = 0 .$$

It is also possible to avoid both the second and third order derivatives conditions in (2.10)-(2.11) by those in Huber (1967) and Inagaki (1973). But, these alternative conditions, in turn, require extra conditions on the first and second order moments of

$$(2.13)\quad \sup_{\underset{\sim}{\theta}:\|\underset{\sim}{\theta}-\underset{\sim}{\theta}_0\|<\delta}\left\|(\partial/\partial\underset{\sim}{\theta})\log f_i(x,\underset{\sim}{\theta})\Big]_{\underset{\sim}{\theta}} - (\partial/\partial\underset{\sim}{\theta})\log f_i(x,\underset{\sim}{\theta})\Big]_{\underset{\sim}{\theta}_0}\right\|$$

for small $\delta(>0)$. In the sequel, we shall deal with (2.10)-(2.11) only — though towards the end of Section 3, we shall make certain comments on these alternative conditions.]

[A4]: For every $i(=1,\dots,k)$ and $\underset{\sim}{\theta}\in\Omega$,

$$(2.14)\quad \int_{E^p}(\partial^2/\partial\theta_j\partial\theta_\ell)f_i(x,\underset{\sim}{\theta})d\mu(\underset{\sim}{x}) = 0 ,\quad \forall\ j,\ell = 1,\dots,t .$$

Let us define for each $i(=1,\dots,k)$

$$\underset{\sim}{B}_{\underset{\sim}{\theta}}^{(i)} = \left(\left(\int_{E^p}(\partial/\partial\theta_j)\log f_i(x,\underset{\sim}{\theta})(\partial/\partial\theta_\ell)\log f_i(x,\underset{\sim}{\theta})dF_i(x,\underset{\sim}{\theta})\right)\right)_{j,\ell=1,\dots,t} .$$
(2.15)

[A5]: $\underset{\sim}{B}_{\underset{\sim}{\theta}}^{(1)},\dots,\underset{\sim}{B}_{\underset{\sim}{\theta}}^{(k)}$ are all continuous in $\underset{\sim}{\theta}$ in some neighbourhood of $\underset{\sim}{\theta}_0$ and

$$(2.16)\quad \underset{\sim}{B}_{\underset{\sim}{\theta}_0}^{*} = \sum_{i=1}^{k}(n_i/n)\underset{\sim}{B}_{\underset{\sim}{\theta}_0}^{(i)} \text{ is positive definite .}$$

[A6]: $\underset{\sim}{h}(\underset{\sim}{\theta})$ posssesses continuous first and second order derivatives with respect to $\underset{\sim}{\theta}$, $\forall\ \underset{\sim}{\theta} \in \Omega$. Let then

$$(2.17)\qquad \underset{\sim}{H}_{\underset{\sim}{\theta}} = (((\partial/\partial\underset{\sim}{\theta})\underset{\sim}{h}(\underset{\sim}{\theta}))) \quad \text{(of order } t \times r) .$$

[A7]: $\underset{\sim}{H}_{\underset{\sim}{\theta}_0}$ is of rank $r(< t)$.

[A8]: The following matrix (of order $(r+t) \times (r+t)$)

$$(2.18)\qquad \begin{pmatrix} \underset{\sim}{B}^*_{\underset{\sim}{\theta}_0} & -\underset{\sim}{H}_{\underset{\sim}{\theta}_0} \\ -\underset{\sim}{H}'_{\underset{\sim}{\theta}_0} & \underset{\sim}{0} \end{pmatrix} \text{ is of full-rank}$$

and we denote by

$$(2.19)\qquad \begin{bmatrix} \underset{\sim}{B}^*_{\underset{\sim}{\theta}_0} & -\underset{\sim}{H}_{\underset{\sim}{\theta}_0} \\ -\underset{\sim}{H}'_{\underset{\sim}{\theta}_0} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \underset{\sim}{P}^*_{\underset{\sim}{\theta}_0} & \underset{\sim}{Q}^*_{\underset{\sim}{\theta}_0} \\ \underset{\sim}{Q}^{*\prime}_{\underset{\sim}{\theta}_0} & \underset{\sim}{R}^*_{\underset{\sim}{\theta}_0} \end{bmatrix} .$$

Note that $\underset{\sim}{B}^*_{\underset{\sim}{\theta}_0}$ (and hence, $\underset{\sim}{P}^*_{\underset{\sim}{\theta}_0}$, $\underset{\sim}{Q}^*_{\underset{\sim}{\theta}_0}$ and $\underset{\sim}{R}^*_{\underset{\sim}{\theta}_0}$) may depend on n through $n_1,\dots,n_k$. We make the final assumption:

[A9]: $\lim_{n\to\infty} \frac{1}{n} n_i = \rho_i$ and $(0 < \rho_i < 1)$ exists, $\forall\ 1 \le i \le k$ and $\sum_{i=1}^k \rho_i = 1$.

Under [A9], $\underset{\sim}{B}^*_{\underset{\sim}{\theta}_0}$ converges to

$$(2.20)\qquad \bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0} = \sum_{i=1}^k \rho_i \underset{\sim}{B}^{(i)}_{\underset{\sim}{\theta}_0}$$

and in (2.19), on replacing $\underset{\sim}{B}^*$ by $\bar{\underset{\sim}{B}}$, the corresponding matrices on the right hand side (rhs) are denoted by $\bar{\underset{\sim}{P}}_{\underset{\sim}{\theta}_0}$, $\bar{\underset{\sim}{Q}}_{\underset{\sim}{\theta}_0}$ and $\bar{\underset{\sim}{R}}_{\underset{\sim}{\theta}_0}$ respectively; these do not depend on n. Note that, by definition,

$$(2.21)\qquad \begin{bmatrix} \bar{\underset{\sim}{P}}_{\underset{\sim}{\theta}_0} & \bar{\underset{\sim}{Q}}_{\underset{\sim}{\theta}_0} \\ \bar{\underset{\sim}{Q}}'_{\underset{\sim}{\theta}_0} & \bar{\underset{\sim}{R}}_{\underset{\sim}{\theta}_0} \end{bmatrix} \begin{bmatrix} \bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0} & -\underset{\sim}{H}_{\underset{\sim}{\theta}_0} \\ -\underset{\sim}{H}'_{\underset{\sim}{\theta}_0} & \underset{\sim}{0} \end{bmatrix} = \begin{pmatrix} \underset{\sim}{I} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{I} \end{pmatrix} .$$

Note that $\bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}$, $\bar{\underset{\sim}{P}}_{\underset{\sim}{\theta}_0}$ and $\bar{\underset{\sim}{R}}_{\underset{\sim}{\theta}_0}$ are all symmetric matrices. For later use, we also define the vectors,

$$(2.22) \qquad \underset{\sim}{\Lambda}_n(\underset{\sim}{\theta}) = n^{-\frac{1}{2}}(\partial/\partial\underset{\sim}{\theta}) \log L_n(\underset{\sim}{X}_n,\underset{\sim}{\theta}) \ , \quad \underset{\sim}{\Lambda}_n^0 = \underset{\sim}{\Lambda}_n(\underset{\sim}{\theta}_0) \ ,$$

and for vectors or matrices use the notations $\underset{\sim}{o}_p$ and $\underset{\sim}{O}_p$ (or $\underset{\sim}{o}$ and $\underset{\sim}{O}$) the sense that these orders apply to the individual elements of them.

2. ASYMPTOTIC DISTRIBUTION THEORY UNDER H_0: $\underset{\sim}{\theta}_0 \in \omega$

First, consider the case of the unrestricted MLE $\tilde{\underset{\sim}{\theta}}_n$. Under the assumptions made in Section 2, $\underset{\sim}{\theta}_n$ exists (a.e.), it almost surely (a.s.) converges to $\underset{\sim}{\theta}_0$ and further [viz., Silvey (1959)], as $n \to \infty$,

$$(3.1) \qquad n^{\frac{1}{2}}(\tilde{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_0) = \bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}^{-1}\underset{\sim}{\Lambda}_n^0 + \underset{\sim}{o}_p(1) \ .$$

Also, by a direct application of the multivariate central limit theorem,

$$(3.2) \qquad \underset{\sim}{\Lambda}_n^0 \overset{\mathcal{D}}{\to} N_t(\underset{\sim}{0},\bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}) \ , \quad \text{as} \quad n \to \infty \ .$$

Consequently, from (3.1) and (3.2), we have

$$(3.3) \qquad n^{\frac{1}{2}}(\tilde{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_0) \overset{\mathcal{D}}{\to} N_t(\underset{\sim}{0},\bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}^{-1}) \quad \text{as} \quad n \to \infty \ .$$

For the restricted MLE $\hat{\underset{\sim}{\theta}}_n$, consider the equations

$$(3.4) \qquad \begin{cases} n^{-\frac{1}{2}}\underset{\sim}{\Lambda}_n(\underset{\sim}{\theta}) + \underset{\sim}{H}_{\underset{\sim}{\theta}}\underset{\sim}{\lambda} = \underset{\sim}{0} \\ \underset{\sim}{h}(\underset{\sim}{\theta}) = 0 \end{cases}$$

where $\lambda(\in E^r)$ is a Lagrangian multiplier vector; the solutions for θ and λ are $\hat{\theta}_n$ and $\hat{\lambda}_n$, respectively. From the results in Section 7 of Silvey (1959), we conclude that under $H_0: \theta_0 \in \omega$ and the regularity conditions of Section 2, as $n \to \infty$,

(3.5) $$n^{\frac{1}{2}}(\hat{\theta}_n - \theta_0) = \bar{P}_{\theta_0} \Lambda_n^0 + o_p(1) ,$$

(3.6) $$n^{\frac{1}{2}}\hat{\lambda}_n = \bar{Q}'_{\theta_0} \Lambda_n^0 + o_p(1)$$

and further for the likelihood ratio statistic L_h in (2.5)

(3.7) $$L_n = n(\hat{\theta}_n - \tilde{\theta}_n)' \bar{B}_{\theta_0} (\hat{\theta}_n - \tilde{\theta}_n) + o_p(1) .$$

Note that by (2.21), $\bar{B}_{\theta_0} \bar{P}_{\theta_0} = I + H_{\theta_0} \bar{Q}'_{\theta_0}$, $\bar{B}_{\theta_0} \bar{Q}_{\theta_0} = H_{\theta_0} \bar{R}_{\theta_0}$, $-H'_{\theta_0} \bar{P}_{\theta_0} = 0$ and $-H'_{\theta_0} \bar{Q}_{\theta_0} = I$. Thus, noting that $\bar{P}'_{\theta_0} = \bar{P}_{\theta_0}$, we have

(3.8) $$\bar{P}_{\theta_0} \bar{B}_{\theta_0} \bar{P}'_{\theta_0} = \bar{P}_{\theta_0} \bar{B}_{\theta_0} \bar{P}_{\theta_0} = \bar{P}_{\theta_0} + \bar{P}_{\theta_0} H_{\theta_0} \bar{Q}'_{\theta_0} = \bar{P}_{\theta_0} + \bar{P}'_{\theta_0} H_{\theta_0} \bar{Q}'_{\theta_0} = \bar{P}_{\theta_0} .$$

Hence, from (3.2), (3.5) and (3.8), we have under $H_0: \theta_0 \in \omega$

(3.9) $$n^{\frac{1}{2}}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{D}} N_t(0, \bar{P}_{\theta_0}) , \quad \text{as } n \to \infty .$$

Also, from (3.1), (3.5), (3.6) and (3.7), we have (on using the identities presented before (3.8)) under $H_0: \theta_0 \in \omega$,

(3.10) $$\begin{aligned} L_n &= \Lambda_n^{0'} (\bar{P}_{\theta_0} - \bar{B}^{-1}_{\theta_0}) \bar{B}_{\theta_0} (\bar{P}_{\theta_0} - \bar{B}^{-1}_{\theta_0}) \Lambda_n^0 + o_p(1) \\ &= \Lambda_n^{0'} \bar{Q}_{\theta_0} H'_{\theta_0} \bar{B}^{-1}_{\theta_0} H_{\theta_0} \bar{Q}'_{\theta_0} \Lambda_n^0 + o_p(1) \\ &= -\hat{\lambda}'_n \bar{R}^{-1}_{\theta_0} \hat{\lambda}_n + o_p(1) \\ &= -\Lambda_n^{0'} \bar{Q}_{\theta_0} \bar{R}^{-1}_{\theta_0} \bar{Q}'_{\theta_0} \Lambda_n^0 + o_p(1) \end{aligned}$$

where it is easy to show (by using (2.21)) that

$$\bar{Q}_{\theta_0} \bar{R}_{\theta_0}^{-1} \bar{Q}'_{\theta_0} \bar{B}_{\theta_0}^{-1} \bar{Q}_{\theta_0} \bar{R}_{\theta_0} \bar{Q}_{\theta_0} = -\bar{Q}_{\theta_0} \bar{R}_{\theta_0}^{-1} \bar{Q}'_{\theta_0} , \tag{3.11}$$

$$\text{Rank of } \left\{-\bar{Q}_{\theta_0} R_{\theta_0}^{-1} \bar{Q}'_{\theta_0}\right\} = r(< t) . \tag{.12}$$

Hence, by (3.2), (3.10), (3.11), (3.12) and the Cochran theorem on quadratic forms in (asymptotically) normally distributed random vectors, we obtain that under H_0: $\theta_0 \in \omega$,

$$L_n \overset{\mathcal{D}}{\to} \chi_r^2 , \tag{3.13}$$

and let $\chi_{r,\alpha}^2$ be the upper $100\alpha\%$ point of the chi-square d.f. with r degrees of freedom (DF). Then, from (2.6) and (3.13), we have

$$\ell_{n,\alpha} \to \chi_{r,\alpha}^2 \quad \text{as} \quad n \to \infty , \tag{3.14}$$

Let us now consider the case of $\{\theta_n^*\}$. By (2.7), we have for every $y \in E^t$,

$$\begin{aligned} P\{n^{\frac{1}{2}}(\theta_n^* - \theta_0) \le y | H_0\} \\ &= P\{n^{\frac{1}{2}}(\hat{\theta}_n - \theta_0) \le y , \quad L_n \le \ell_{n,\alpha} | H_0\} \\ &\quad + P\{n^{\frac{1}{2}}(\tilde{\theta}_n - \theta_0) \le y , \quad L_n > \ell_{n,\alpha} | H_0\} . \end{aligned} \tag{3.15}$$

By (3.2), (3.5), (3.6), (3.10) and (3.14), the first term on the rhs of (3.15) converges to

$$P\{\bar{P}_{\theta_0} \Lambda_n^0 \le y , \quad -n\hat{\lambda}_n' \bar{R}_{\theta_0}^{-1} \hat{\lambda}_n \le \chi_{r,\alpha}^2 | H_0\} . \tag{3.16}$$

Note that by (2.21), $\bar{Q}'_{\theta_0}\bar{B}_{\theta_0}\bar{P}_{\theta_0} = \bar{Q}'_{\theta_0} + \bar{Q}'_{\theta_0}H_{\theta_0}\bar{Q}_{\theta_0} = \bar{Q}_{\theta_0} - \bar{Q}_{\theta_0} = 0$, so that $\bar{P}_{\theta_0}\Lambda_n^0$ and $n^{\frac{1}{2}}\hat{\lambda}_n$ are asymptotically independent, and hence, (3.16) reduces (asymptotically) to

$$P\{\bar{P}_{\theta_0}\Lambda_n^0 \le y \mid H_0\}P\{L_n \le \chi^2_{r,\alpha} \mid H_0\} \tag{3.17}$$

$$\to (1-\alpha)G_t(y;\ 0,\ \bar{P}_{\theta_0})\ ,$$

where $G_t(y;\ \mu,\ \Sigma)$ stands for a t-variate multinormal df with mean vector μ and dispersion matrix Σ. Let us also denote by

$$E_1 = \{x \in E^r:\ -x'\bar{R}^{-1}_{\theta_0}x > \chi^2_{r,\alpha}\}\ . \tag{3.18}$$

Then, by (3.1), (3.10) and (3.14), the second term on the rhs of (3.15) is asymptotically equivalent to

$$P\{\bar{B}^{-1}_{\theta_0}\Lambda_n^0 \le y\ ,\ \ n^{\frac{1}{2}}\hat{\lambda}_n \in E_1 \mid H_0\} \tag{3.19}$$

where by (3.2) and (3.6), under H_0, $(\bar{B}^{-1}_{\theta_0}\Lambda_n^0,\ n^{\frac{1}{2}}\hat{\lambda}_n)$ has asymptotically a(t+r)-variate normal df with 0 mean and dispersion matrix

$$\begin{pmatrix} \bar{B}^{-1}_{\theta_0} & \bar{Q}_{\theta_0} \\ \bar{Q}'_{\theta_0} & -R_{\theta_0} \end{pmatrix}, \tag{3.20}$$

so that the conditional df of $\bar{B}^{-1}_{\theta_0}\Lambda_n^0$ given $n^{\frac{1}{2}}\hat{\lambda}_n = z$ is asymptotically multi-normal with mean vector $\bar{Q}_{\theta_0}\bar{R}_{\theta_0}z$ and dispersion matrix $\bar{B}^{-1}_{\theta_0}$ + $\bar{Q}_{\theta_0}\bar{R}_{\theta_0}\bar{Q}'_{\theta_0} = \bar{P}_{\theta_0}$. Hence, (3.19) converges to

$$(3.21)\qquad \int_{E_1} G_t(\underset{\sim}{y}-\underset{\sim}{\overline{Q}}_{\underset{\sim}{\theta}_0}\underset{\sim}{\overline{R}}^{-1}_{\underset{\sim}{\theta}_0}\underset{\sim}{Z};\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\underset{\sim}{\theta}_0})dG_r(\underset{\sim}{Z};\underset{\sim}{0},-\underset{\sim}{\overline{R}}_{\underset{\sim}{\theta}_0}) .$$

From (3.15), (317), and (3.21), we obtain that under H_0: $\underset{\sim}{\theta}_0 \in \omega$,

$$(3.22)\quad \lim_{n\to\infty} P\{n^{\frac{1}{2}}(\underset{\sim}{\theta}^*_n - \underset{\sim}{\theta}_0) \le \underset{\sim}{y}\}$$

$$= (1-\alpha)G_t(\underset{\sim}{y};\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\underset{\sim}{\theta}_0})$$

$$+ \int_{E_1} G_t(\underset{\sim}{y}+\underset{\sim}{\overline{Q}}_{\underset{\sim}{\theta}_0}\underset{\sim}{\overline{R}}^{-1}_{\underset{\sim}{\theta}_0}\underset{\sim}{Z};\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\underset{\sim}{\theta}_0})dG_r(\underset{\sim}{Z};\underset{\sim}{0},-\underset{\sim}{\overline{R}}_{\underset{\sim}{\theta}_0}) , \quad \forall\ \underset{\sim}{y} \in E^t .$$

Thus, the asymptotic distribution of $n^{\frac{1}{2}}(\underset{\sim}{\theta}^*_n - \theta_0)$ is, in general, non-normal. In particular, if $\underset{\sim}{\overline{Q}}_{\underset{\sim}{\theta}_0}\underset{\sim}{\overline{R}}^{-1}_{\underset{\sim}{\theta}_0} = \underset{\sim}{0}$ (which also implies that $\underset{\sim}{\overline{P}}_{\underset{\sim}{\theta}_0} = \underset{\sim}{\overline{B}}^{-1}_{\underset{\sim}{\theta}_0}$), (3.22) reduces to

$$(3.23)\qquad G_t(\underset{\sim}{y};\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\underset{\sim}{\theta}_0}) \quad (\text{or}\quad G_t(\underset{\sim}{y};\underset{\sim}{0},\underset{\sim}{\overline{B}}^{-1}_{\underset{\sim}{\theta}_0}))$$

so that all the three stimators have the same limiting normal distribution.

Remark. Our (3.1), (3.5) and (3.6), as adapted from Silvey (1959), rest on the assumptions made in Section 2. As mentioned in Section 2, (2.12) may replace (2.10)-(2.11) for $s = 3$. In such a case, our (3.1), (3.5) and (3.6), would follow from the results of Feder (1968). Also, we may proceed as in Inagaki (1973) and show that (3.1), (3.5) and (3.6) follow, provided (2.13) satisfies appropriate growth conditions. Rest of the formulae in this section remains the same irrespective to the particular approach we choose.

4. ASYMPTOTIC NON-NULL DISTRIBUTION THEORY

Note that if H_0: $\underset{\sim}{\theta}_0 \in \omega$ does not hold (i.e., $\underset{\sim}{h}(\underset{\sim}{\theta}_0) \neq \underset{\sim}{0}$), then there exists a $\underset{\sim}{\theta}^*(\in \omega)$ such that

$$(4.1) \qquad (\partial/\partial\underset{\sim}{\theta})Z(\underset{\sim}{\theta}) + \underset{\sim}{H}_{\underset{\sim}{\theta}}\underset{\sim}{\lambda}^*\big|_{\underset{\sim}{\theta}=\underset{\sim}{\theta}^*} = \underset{\sim}{0} \quad \text{and} \quad \underset{\sim}{h}(\underset{\sim}{\theta}^*) = \underset{\sim}{0}$$

where $\underset{\sim}{\lambda}^*(\in E^r)$ is a Lagrangian multiplier and $\underset{\sim}{\theta}^* \neq \underset{\sim}{\theta}_0$ (by assumptions [A1] and [A2]). In this case, $\tilde{\underset{\sim}{\theta}}_n$ in (2.2), stochastically converges to $\underset{\sim}{\theta}_0$ while $\hat{\underset{\sim}{\theta}}_n$ in (2.4) converges stochastically to $\underset{\sim}{\theta}^*$, and hence, L_n in (2.5) tends (in probability) to ∞ as $n \to \infty$. Consequently, by (2.7) and (3.14),

$$(4.2) \qquad \lim_{n\to\infty} P\{\underset{\sim}{\theta}_n^* \neq \underset{\sim}{\theta}_n \mid \underset{\sim}{\theta}_0 \notin \omega\} \leq \lim_{n\to\infty} P\{L_n \leq \ell_{n,\alpha} \mid \underset{\sim}{\theta}_0 \notin \omega\} = 0 \ ,$$

and hence, noting that (3.3) does not depend on H_0: $\underset{\sim}{\theta}_0 \in \omega$ being true or not, we have from (3.3) and (4.2), for every $\underset{\sim}{y} \in E^t$,

$$(4.3) \qquad \lim_{n\to\infty} P\{n^{\frac{1}{2}}(\underset{\sim}{\theta}_n^* - \underset{\sim}{\theta}_0) \leq \underset{\sim}{y} \mid \underset{\sim}{\theta}_0 \notin \omega\}$$
$$= \lim_{n\to\infty} P\{n^{\frac{1}{2}}(\tilde{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_0) \leq \underset{\sim}{y} \mid \underset{\sim}{\theta}_0 \notin \omega\} = G_t(\underset{\sim}{y};\underset{\sim}{0},\overline{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}^{-1}) \ .$$

Thus, for any (fixed) alternative, $n^{\frac{1}{2}}(\underset{\sim}{\theta}_n^* - \underset{\sim}{\theta}_0)$ and $n^{\frac{1}{2}}(\tilde{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_0)$ are asymptotically equivalent in probability and have the same (asymptotic) multinormal distribution. The situation becomes different when $\underset{\sim}{\theta}_0$ lies *near* the boundary of ω. For this study, we conceive of the following sequence $\{K_n\}$ of local alternatives:

$$(4.4) \qquad K_n: \underset{\sim}{h}(\underset{\sim}{\theta}_0) = n^{-\frac{1}{2}}\underset{\sim}{\gamma} \ , \quad \underset{\sim}{\gamma} \ \text{real-vector} \quad (\in E^r) \ ,$$

and consider the asympttoic distributions of the estimators under $\{K_n\}$.

First, (3.3) holds irrespective of H_0 or $\{K_n\}$, and hence

$$\lim_{n\to\infty} P\{n^{\frac{1}{2}}(\tilde{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_0) \le \underset{\sim}{y} \mid K_n\} = G_t(\underset{\sim}{y}; \underset{\sim}{0}, \overline{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}^{-1}) \ , \quad \forall\ \underset{\sim}{y} \in E^t \ . \tag{4.5}$$

Also, under K_n, the solutions $\underset{\sim}{\theta}^*$, $\underset{\sim}{\lambda}^*$ in (4.1) depends on n and are denoted by $\underset{\sim}{\theta}^*_{(n)}$, $\underset{\sim}{\lambda}^*_{(n)}$, respectively. Note that $\underset{\sim}{h}(\underset{\sim}{\theta}^*_{(n)}) = \underset{\sim}{0}$. Hence, under (4.4) and the assumptions of Section 2, we have

$$\underset{\sim}{H}'_{\underset{\sim}{\theta}_0}\{n^{\frac{1}{2}}(\underset{\sim}{\theta}^*_{(n)} - \underset{\sim}{\theta}_0)\} = \underset{\sim}{\gamma} + \underset{\sim}{o}(1) \ , \tag{4.6}$$

$$\underset{\sim}{H}_{\underset{\sim}{\theta}_0}\{n^{\frac{1}{2}}\underset{\sim}{\lambda}^*_{(n)}\} = \overline{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}\{n^{\frac{1}{2}}(\underset{\sim}{\theta}^*_{(n)} - \underset{\sim}{\theta}_0)\} + \underset{\sim}{o}(1) \ . \tag{4.7}$$

From (4.6) and (4.7), we conclude that

$$\lim_{n\to\infty} n^{\frac{1}{2}}(\underset{\sim}{\theta}^*_{(n)} - \underset{\sim}{\theta}_0) = \underset{\sim}{\gamma}^* \quad \text{and} \quad \lim_{n\to\infty} n^{\frac{1}{2}}\underset{\sim}{\lambda}^*_{(n)} = \underset{\sim}{\lambda}^* \tag{4.8}$$

both exist, and

$$\underset{\sim}{\gamma} = \underset{\sim}{H}'_{\underset{\sim}{\theta}_0}\underset{\sim}{\gamma}^*,\ \underset{\sim}{H}_{\underset{\sim}{\theta}_0}\underset{\sim}{\lambda}^* = \overline{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}\underset{\sim}{\gamma}^* \quad (\Longrightarrow \underset{\sim}{\lambda}^* = -\overline{\underset{\sim}{R}}'_{\underset{\sim}{\theta}_0}\overline{\underset{\sim}{H}}'_{\underset{\sim}{\theta}_0}\underset{\sim}{\gamma}^*,\ \underset{\sim}{\gamma}^* = \overline{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}^{-1}\underset{\sim}{H}_{\underset{\sim}{\theta}_0}\underset{\sim}{\lambda}^*) \ . \tag{4.9}$$

From (3.4), (4.4), (4.6) through (4.9) and under the assumptions of Section 2, we obtain that under $\{K_n\}$,

$$n^{\frac{1}{2}}(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_0) = \underset{\sim}{\gamma}^* + \overline{\underset{\sim}{P}}_{\underset{\sim}{\theta}_0}\underset{\sim}{\Lambda}^0_n + \underset{\sim}{o}_p(1) \ , \tag{4.10}$$

$$n^{\frac{1}{2}}\hat{\underset{\sim}{\lambda}}_n = \underset{\sim}{\lambda}^* + \overline{\underset{\sim}{Q}}'_{\underset{\sim}{\theta}_0}\underset{\sim}{\Lambda}^0_n + \underset{\sim}{o}_p(1) \ , \tag{4.11}$$

where $\hat{\underset{\sim}{\theta}}_n$ is the restricted MLE and $\hat{\underset{\sim}{\lambda}}_n$ is the lagrangian multiplier in (3.4). Comparing (3.5) and (4.10) and using the same arugments as in (3.8) and (3.9), we obtain that for every $\underset{\sim}{y} \in E^t$,

$$\lim_{n\to\infty} P\{n^{\frac{1}{2}}(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_0) \le \underset{\sim}{y} \mid \underset{\sim}{K}_n\} = G_t(\underset{\sim}{y} - \underset{\sim}{\gamma}^*; \underset{\sim}{0}, \overline{\underset{\sim}{P}}_{\underset{\sim}{\theta}_0}) \ . \tag{4.12}$$

It follows similarly that (3.7) continues to hold under $\{K_n\}$, where by (3.1) and (4.10), we have

$$(4.13)\qquad n^{\frac{1}{2}}(\hat{\underset{\sim}{\theta}}_n - \tilde{\underset{\sim}{\theta}}_n) = \underset{\sim}{\gamma}^* + (\bar{\underset{\sim}{P}}_{\underset{\sim}{\theta}_0} - \bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}^{-1})\underset{\sim}{\Lambda}_n^0 + \underset{\sim}{o}_p(1)$$

$$= \bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}^{-1}\underset{\sim}{H}_{\underset{\sim}{\theta}_0}\underset{\sim}{\lambda}^* + \bar{\underset{\sim}{Q}}_{\underset{\sim}{\theta}_0}\underset{\sim}{H}'_{\underset{\sim}{\theta}_0}\bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}^{-1}\underset{\sim}{\Lambda}_n^0 + \underset{\sim}{o}_p(1)$$

$$= \bar{\underset{\sim}{Q}}_{\underset{\sim}{\theta}_0}\bar{\underset{\sim}{R}}_{\underset{\sim}{\theta}_0}^{-1}\{n^{\frac{1}{2}}\hat{\underset{\sim}{\lambda}}_n\} + \underset{\sim}{o}_p(1) .$$

Thus, under $\{K_n\}$, as $n \to \infty$

$$(4.14)\qquad L_n = -n\hat{\underset{\sim}{\lambda}}_n \bar{\underset{\sim}{R}}_{\underset{\sim}{\theta}_0}^{-1}\hat{\underset{\sim}{\lambda}}_n + o_p(1) .$$

On noting that, by (2.21), $\bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}\bar{\underset{\sim}{Q}}_{\underset{\sim}{\theta}_0} = \underset{\sim}{H}_{\underset{\sim}{\theta}_0}\bar{\underset{\sim}{R}}_{\underset{\sim}{\theta}_0}$ and $\underset{\sim}{H}_{\underset{\sim}{\theta}_0}\bar{\underset{\sim}{Q}}_{\underset{\sim}{\theta}_0} = -\underset{\sim}{I}$, we conclude from (4.11) and (4.14) that under $\{K_n\}$, L_n has asymptotically a non-central chi-square df with r DF and non-centrality parameter

$$(4.15)\qquad \Delta^* = \underset{\sim}{\lambda}^{*\prime}\underset{\sim}{H}'_{\underset{\sim}{\theta}_0}\bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}^{-1}\underset{\sim}{H}_{\underset{\sim}{\theta}_0}\underset{\sim}{\lambda}^* = -\underset{\sim}{\lambda}^{*\prime}\bar{\underset{\sim}{R}}_{\underset{\sim}{\theta}_0}^{-1}\underset{\sim}{\lambda}^* ;$$

we denote this df by $\Pi_r(x;\Delta^*)$. Then, from (2.7) and (3.14),

$$(4.16)\qquad P\{n^{\frac{1}{2}}(\underset{\sim}{\theta}_n^* - \underset{\sim}{\theta}_0) \le \underset{\sim}{y} | K_n\}$$

$$= P\{n^{\frac{1}{2}}(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_0) \le \underset{\sim}{y},\ L_n \le \ell_{n,\alpha} | K_n\}$$

$$+ P\{n^{\frac{1}{2}}(\tilde{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_0) \le \underset{\sim}{y},\ L_n > \ell_{n,\alpha} | K_n\} ,\quad \forall\ \underset{\sim}{y} \in E^t .$$

Since by (2.21), $\bar{\underset{\sim}{P}}'_{\underset{\sim}{\theta}_0}\bar{\underset{\sim}{B}}_{\underset{\sim}{\theta}_0}\bar{\underset{\sim}{Q}}_{\underset{\sim}{\theta}_0} = \bar{\underset{\sim}{P}}'_{\underset{\sim}{\theta}_0}\underset{\sim}{H}_{\underset{\sim}{\theta}_0}\bar{\underset{\sim}{R}}_{\underset{\sim}{\theta}_0} = \underset{\sim}{0}$, we conclude from (4.10) and (4.11) that under K_n, $n^{\frac{1}{2}}(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_0)$ and $n^{\frac{1}{2}}\hat{\underset{\sim}{\lambda}}_n$ (and hence, by (4.14), L_n) are asymptotically independent, so that by (3.2), (4.10), (4.11), (4.14) and (4.15), the first term on the rhs of (4.16) converges to

$$G_t(\underset{\sim}{y}-\underset{\sim}{\gamma}^*;\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\theta_0})\Pi_r(\chi^2_{r,\alpha};\Delta^*) \quad . \tag{4.17}$$

Also, let

$$E^*(\underset{\sim}{c}) = \{\underset{\sim}{y}\in E^r: -(\underset{\sim}{y}+\underset{\sim}{c})'\underset{\sim}{\overline{R}}^{-1}_{\theta_0}(\underset{\sim}{y}+\underset{\sim}{c}) > \chi^2_{r,\alpha}\} \quad , \quad \forall\ \underset{\sim}{c}\in E^r \ . \tag{4.18}$$

Then, by (3.1), (3.2), (4.11), (4.14) [and the fact that by (2.21), $\underset{\sim}{\overline{Q}}_{\theta_0}\underset{\sim}{\overline{R}}^{-1}_{\theta_0}\underset{\sim}{Q}'_{\theta_0} = \underset{\sim}{\overline{P}}_{\theta_0} - \underset{\sim}{\overline{B}}^{-1}_{\theta_0}$], we obtain that the second term on the rhs of (4.16) is

$$\begin{aligned}(4.19)\qquad & P\{\underset{\sim}{\overline{B}}^{-1}_{\theta_0}\underset{\sim}{\Lambda}^0_n \le \underset{\sim}{y},\ n^{\frac{1}{2}}\hat{\underset{\sim}{\lambda}}_n \in E_1(\underset{\sim}{0})\,|\,K_n\} \\ &= \int_{E_1(\underset{\sim}{0})} G_t(\underset{\sim}{y}+\underset{\sim}{Q}_{\theta_0}\underset{\sim}{\overline{R}}^{-1}_{\theta_0}(\underset{\sim}{z}-\lambda^*);\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\theta_0})\,dG_r(\underset{\sim}{z};\underset{\sim}{\lambda}^*,-\underset{\sim}{\overline{R}}_{\theta_0}) + o(1) \\ &= \int_{E_1(\underset{\sim}{\lambda}^*)} G_t(\underset{\sim}{y}+\underset{\sim}{\overline{Q}}_{\theta_0}\underset{\sim}{\overline{R}}^{-1}_{\theta_0}\underset{\sim}{x};\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\theta_0})\,dG_r(\underset{\sim}{x};\underset{\sim}{0},-\underset{\sim}{\overline{R}}_{\theta_0}) + o(1) \ .\end{aligned}$$

From (4.16), (4.17) and (4.19), we arrive at the following.

<u>Theorem 4.1</u>. *Under* $\{K_n\}$ *in* (4.4) *and the assumptions of Section* 2,

$$\begin{aligned}(4.20)\qquad & \lim_{n\to\infty} P\{n^{\frac{1}{2}}(\underset{\sim}{\theta}^*_n-\underset{\sim}{\theta}_0) \le \underset{\sim}{y}\} = G_t(\underset{\sim}{y}-\underset{\sim}{\gamma}^*;\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\theta_0})\Pi_r(\chi^2_{r,\alpha};\Delta^*) \\ &\quad + \int_{E_1(\underset{\sim}{\lambda}^*)} G_t(\underset{\sim}{y}+\underset{\sim}{\overline{Q}}_{\theta_0}\underset{\sim}{\overline{R}}^{-1}_{\theta_0}\underset{\sim}{x};\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\theta_0})\,dG_r(\underset{\sim}{x};\underset{\sim}{0},-\underset{\sim}{\overline{R}}_{\theta_0}) \\ &\quad = G^*_t(\underset{\sim}{y};\underset{\sim}{\gamma}) \ , \ \text{say}, \ (\underset{\sim}{y}\in E^t) \ .\end{aligned}$$

Here also, if $\underset{\sim}{\overline{Q}}_{\theta_0}\underset{\sim}{\overline{R}}^{-1}_{\theta_0} = \underset{\sim}{0}$, (4.20) reduces to

$$G_t(\underset{\sim}{y}-\underset{\sim}{\gamma}^*;\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\theta_0})\Pi_r(\chi^2_{p,\alpha};\Delta^*) + [1-\Pi_r(\chi^2_{p,\alpha};\Delta^*)]G_t(\underset{\sim}{y};\underset{\sim}{0},\underset{\sim}{\overline{P}}_{\theta_0}) \tag{4.21}$$

(that is a mixture of two multinormal df's). But, in general, it is non-normal. For later use, we denote the probability density functions (pdf) corresponding to G_t and G^*_t by g_t and g^*_t, respectively. Then,

$$(4.21)\quad g_t^*(\underset{\sim}{y},\underset{\sim}{\gamma}) = g_t(\underset{\sim}{y}-\underset{\sim}{\gamma}^*;\underset{\sim}{0},\overline{\underset{\sim}{P}}_{\underset{\sim}{\theta}_0})\Pi_r(\chi^2_{r,\alpha};\Delta^*)$$

$$+\int_{E_1(\underset{\sim}{\lambda}^*)} g_t(\underset{\sim}{y}+\overline{\underset{\sim}{Q}}_{\theta_0}\overline{\underset{\sim}{R}}_{\theta_0}^{-1}\underset{\sim}{x};\underset{\sim}{0},\overline{\underset{\sim}{P}}_{\theta_0})g_r(\underset{\sim}{x};\underset{\sim}{0},-\overline{\underset{\sim}{R}}_{\theta_0})d\underset{\sim}{x}\ ,\quad \forall\ \underset{\sim}{y}\in E^t\ .$$

5. ASYMPTOTIC COMPARISON OF THE ESTIMATORS

Let $\{T_n\}$ be a sequence of estimators of $\underset{\sim}{\theta}_0$ such that $n^{\frac{1}{2}}(\underset{\sim}{T}_n-\underset{\sim}{\theta}_0)$ has a limiting distribution with finite second moments. Then the *mean vector* and *dispersion matrix* of this limiting df are taken as the *asymptotic bias* and *asymptotic dispersion matrix* (*a.d.m.*) of $n^{\frac{1}{2}}(\underset{\sim}{T}_n-\underset{\sim}{\theta}_0)$. Here, we study the asymptotic bias and a.d.m. of each of the three estimators considered in earlier sections and compare them. We confine ourselves to the sequence $\{K_n\}$ of alternative hypotheses in (4.4), so that the null hypothesis case follows by letting $\underset{\sim}{\gamma}=\underset{\sim}{0}$.

It follows from (3.1), (3.2) and (3.3) that

$$(5.1)\quad \underset{\sim}{\beta}_1(\underset{\sim}{\gamma}) = \text{Asymptotic bias of } n^{\frac{1}{2}}(\tilde{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta}_0) \text{ when } \{K_n\} \text{ holds}$$

$$= \int_{E^t} \underset{\sim}{y}dG_t(\underset{\sim}{y};\underset{\sim}{0},\overline{\underset{\sim}{B}}_{\theta_0}^{-1}) = \underset{\sim}{0}\ ;$$

$$(5.2)\quad \underset{\sim}{\nu}_1(\underset{\sim}{\gamma}) = \text{A.d.m. of } n^{\frac{1}{2}}(\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta}_0) \text{ when } \{K_n\} \text{ holds}$$

$$= \int_{E^t} \underset{\sim}{y}\underset{\sim}{y}'dG_t(\underset{\sim}{y};\underset{\sim}{0},\overline{\underset{\sim}{B}}_{\theta_0}^{-1}) = \overline{\underset{\sim}{B}}_{\theta_0}^{-1}\ .$$

Similary, from (3.2) and (4.10), we have

$$(5.3)\quad \underset{\sim}{\beta}_2(\underset{\sim}{\gamma}) = \text{Asymptotic bias of } n^{\frac{1}{2}}(\hat{\underset{\sim}{\theta}}_n-\underset{\sim}{\theta}_0) \text{ when } \{K_n\} \text{ holds}$$

$$= \underset{\sim}{\gamma}^* = \overline{\underset{\sim}{Q}}_{\theta_0}\overline{\underset{\sim}{R}}_{\theta_0}^{-1}\underset{\sim}{\lambda}^*$$

(5.4) $$\nu_2(\gamma) = \text{A.d.m. of } n^{\frac{1}{2}}(\hat{\theta}_n - \theta_0) \text{ when } \{K_n\} \text{ holds}$$

$$= \gamma^*\gamma^{*\prime} + \bar{P}_{\theta_0}\bar{B}_{\theta_0}\bar{P}'_{\theta_0} = \gamma^*\gamma^{*\prime} + \bar{P}_{\theta_0} .$$

At this stage, we note that for a multinormal df $G_p(x;0,D)$,

(5.5) $$\int_{(x+a)'D^{-1}(x+a)>c} x\,dG_p(x;0,D) = a[\Pi_p(c,\delta) - \Pi_{p+2}(c,\delta)] ,$$
$$\forall\ a \in E^p ,\quad c \geq 0 ;$$

$$\int_{(x+a)'D^{-1}(x+a)>c} xx'\,dG_p(x;0,D) = \{1 - \Pi_{p+2}(c;\delta)\}D -$$

(5.6) $$aa'\{\Pi_p(c;\delta) - 2\Pi_{p+2}(c'\delta) + \Pi_{p+4}(c;\delta)\}$$

where $\delta = a'D^{-1}a$. Thus, from (4.20), (4.21) and (5.5), we have

(5.7) $$\beta^*(\gamma) = \text{Asymptotic bias of } n^{\frac{1}{2}}(\theta_n^* - \theta_0) \text{ when } \{K_n\} \text{ holds}$$

$$= \gamma^*\Pi_r(\chi^2_{r,\alpha};\Delta^*) - \bar{Q}_{\theta_0}\bar{R}^{-1}_{\theta_0}\lambda^*[\Pi_r(\chi^2_{r,\alpha};\Delta^*) - \Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*)]$$

$$= \bar{Q}_{\theta_0}\bar{R}^{-1}_{\theta_0}\lambda^*\Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*) ,\quad (\text{as } \gamma^* = \bar{Q}_{\theta_0}\bar{R}^{-1}_{\theta_0}\lambda^*) ;$$

from (4.20), (4.21) and (5.6), we have

(5.8) $$\nu^*(\gamma) = \text{A.d.m. of } n^{\frac{1}{2}}(\theta_n^* - \theta_0) \text{ when } \{K_n\} \text{ holds}$$

$$= (\gamma^*\gamma^* + \bar{P}_{\theta_0})\Pi_r(\chi^2_{r,\alpha};\Delta^*) + \bar{P}_{\theta_0}\{1 - \Pi_r(\chi^2_{r,\alpha};\Delta^*)\}$$

$$+ \bar{Q}_{\theta_0}\bar{R}^{-1}_{\theta_0}\{-[1 - \Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*]\bar{R}_{\theta_0} - \lambda^*\lambda^{*\prime}[\Pi_r(\chi^2_{r,\alpha};\Delta^*)$$

$$- 2\Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*) + \Pi_{r+4}(\chi^2_{r,\alpha};\Delta^*)]\}\bar{R}^{-1}_{\theta_0}\bar{Q}'_{\theta_0}$$

$$= \bar{\mathbf{P}}_{\theta_0} - [1 - \Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*)]\bar{\mathbf{Q}}_{\theta_0}\bar{\mathbf{R}}^{-1}_{\theta_0}\bar{\mathbf{Q}}'_{\theta_0}$$

$$+ \gamma^*\gamma^{*\prime}\{2\Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*) - \Pi_{r+4}(\chi^2_{r,\alpha};\Delta^*)\}$$

$$= \bar{\mathbf{B}}^{-1}_{\theta_0} + (\bar{\mathbf{P}}_{\theta_0} - \bar{\mathbf{B}}^{-1}_{\theta_0})\Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*) + \gamma^*\gamma^{*\prime}\{2\Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*) - \Pi_{r+4}(\chi^2_{r,\alpha};\Delta^*)\} .$$

We now proceed to compare the asymptotic bias and a.d.m. of the three estimators. First, concsider the null hypothesis case where $\theta_0 \in \omega$, so that $\gamma^* = 0$, $\gamma = 0$, $\lambda^* = 0$. From (5.1), (5.3) and (5.7), we have

(5.9) $$\gamma^* = 0 \Rightarrow \beta_1(0) = \beta_2(0) = \beta^*(0) = 0 ,$$

so that all these estimators are asymptotically unbiased. Also, from (5.2), (5.4) and (5.8), we have for $\gamma = 0$,

(5.10) $$\nu_1(0) = \bar{\mathbf{B}}^{-1}_{\theta_0},\ \nu_2(0) = \bar{\mathbf{P}}_{\theta_0} \quad \text{and} \quad \nu^*(0) = \bar{\mathbf{B}}^{-1}_{\theta_0} - (\bar{\mathbf{B}}^{-1}_{\theta_0} - \bar{\mathbf{P}}_{\theta_0})\Pi_{r+2}(\chi^2_{r,\alpha};0)$$

where

(5.11) $$0 < \Pi_{r+2}(\chi^2_{r,\alpha};0) < \Pi_r(\chi^2_{r,\alpha};0) = 1 - \alpha < 1 .$$

Also, by the identities in (3.8) and prior to it, we note that both $\bar{\mathbf{B}}_{\theta_0}\bar{\mathbf{P}}_{\theta_0}$ and $-\mathbf{H}_{\theta_0}\bar{\mathbf{Q}}'_{\theta_0}$ are idempotent matrices and $\bar{\mathbf{B}}_{\theta_0}$ is non-singuler. Hence, it follows that

(5.12) $$\bar{\mathbf{B}}^{-1}_{\theta_0} - \bar{\mathbf{P}}_{\theta_0} \quad \text{is positive semi-definite (p.s.d.)} .$$

From (5.10), (5.11) and (5.12), we conclude that

(5.13) $\nu_1(0) - \nu_2(0)$, $\nu_1(0) - \nu^*(0)$ and $\nu^*(0) - \nu_2(0)$ are all p.s.d.

In the multiparameter case, the relative efficiency may be judged by the generalized variance (D-optimaltiy) or the trace of the covariance matrix (A-optimality) criterion. In view of the fact that by (2.18)-(2.20), both $\bar{B}_{\theta_0}$ and $\bar{P}_{\theta_0}$ are of full-rank, we have $\nu_1(0)$, $\nu_2(0)$ and $\nu^*(0)$ also of full rank. Hence, we have to difficulty in applying the first criterion. Similar results hold for the second criterion too. We define the "asymptotic generalized variance" as the t-th root of the determinant of the a.d.m. In this light, the asymptotic relative efficiency (A.R.E.) of $\{\theta_n^*\}$ with respect to $\{\tilde{\theta}_n\}$ when $H_0: \theta_0 \in \omega$ holds is

$$(5.14) \qquad e_0(\theta^*, \tilde{\theta}) = \{|\nu_1(0)|/|\nu^*(0)|\}^{1/t}$$

$$\geq 1 \ , \quad \text{by (5.13)} \ ,$$

where the equality sign holds when $\bar{B}_{\theta_0}^{-1} = \bar{P}_{\theta_0}$ i.e., $H_{\theta_0} \bar{Q}_{\theta_0}' = 0$ or equivalently, $\bar{Q}_{\theta_0} \bar{R}_{\theta_0}^{-1} = 0$. Similarly, the A.R.E. of $\{\theta_n^*\}$ with respect to $\{\hat{\theta}_n\}$ is

$$(5.15) \qquad e_0(\theta^*, \hat{\theta}) = \{|\nu_2(0)|/|\nu^*(0)|\}^{1/t}$$

$$\leq 1 \ , \quad \text{by (5.13)} \ ,$$

where the equality sign holds when $\bar{Q}_{\theta_0} \bar{R}_{\theta_0}^{-1} = 0$. Thus, under H_0, $\{\theta_n^*\}$ may not perform as well as $\{\hat{\theta}_n\}$; nevertheless, it is asymptotically, at least as good as $\{\tilde{\theta}_n\}$. In this sense, we have the ordered relation

$$(5.16) \qquad \{\tilde{\theta}_n\} < \{\theta_n^*\} < \{\hat{\theta}_n\} \quad \text{when } H_0 \text{ holds} \ .$$

Let us now consider the general case when $\{K_n\}$ holds. It follows from (5.1), (5.3) and (5.7) that

$$(5.17)\qquad \beta_1(\gamma) = 0,\ \beta_2(\gamma) = \gamma^* \text{ and } \beta^*(\gamma) = \Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*)\beta_2(\gamma)\ ,$$

where $0 < \Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*) < \Pi_{r+2}(\chi^2_{r,\alpha};0) < \Pi_r(\chi^2_{r,\alpha};0) = 1-\alpha < 1$. Thus, on noting that

$$(5.18)\qquad \lim_{\delta\uparrow\infty} \delta\Pi_{r+2}(\chi^2_{r,\alpha};\delta) = 0\ ,$$

we conclude that whereas $\tilde{\theta}_n$ is asymptotically unbiased, $\hat{\theta}_n$ and θ^*_n are not so, and, moreover, θ^*_n has smaller asymptotic bias than $\hat{\theta}_n$. Also, the asymptotic bias of $\hat{\theta}_n$ goes to ∞ if $||\gamma|| \to \infty$, where as $||\gamma|| \to \infty$ ($\Rightarrow \Delta^* \to \infty$), the asymptotic bias of $\theta^*_n \to 0$. Thus, θ^*_n has an edge over $\hat{\theta}_n$ with respect to the asymptotic bias. From (5.2) and (5.4), the A.R.E. of $\{\tilde{\theta}_n\}$ with respect to $\{\hat{\theta}_n\}$ is given by

$$(5.19)\qquad \begin{aligned} e(\tilde{\theta},\hat{\theta}|\gamma^*) &= \{|P_{\theta_0} + \gamma^*\gamma^{*\prime}|/|\bar{B}^{-1}_{\theta_0}|\}^{1/t} \\ &= \{|\bar{B}_{\theta_0}\bar{P}_{\theta_0}||I + \bar{P}^{-1}_{\theta_0}\gamma^*\gamma^{*\prime}|\}^{1/t} \\ &= e(\tilde{\theta},\hat{\theta}|0)\{|I + \bar{P}^{-1}_{\theta_0}\gamma^*\gamma^{*\prime}|\}^{1/t} \\ &= e(\tilde{\theta},\hat{\theta}|0)\{1 + \gamma^{*\prime}\bar{P}^{-1}_{\theta_0}\gamma^*\}^{1/t}\ , \end{aligned}$$

as $\gamma^*\gamma^{*\prime}$ is of rank 1, so that $|I + P^{-1}_{\theta_0}\gamma^*\gamma^{*\prime}|$ = product of the characteristic roots of $I + P^{-1}_{\theta_0}\gamma^*\gamma^{*\prime}$ = 1 + largest root of $\bar{P}^{-1}_{\theta_0}\gamma^*\gamma^{*\prime}$ = 1 + $\gamma^{*\prime}\bar{P}^{-1}_{\theta_0}\gamma^*$. Thus, if we let

$$(5.20)\qquad S_1 = \{\gamma^*:\ 1 + \gamma^{*\prime}\bar{P}^{-1}_{\theta_0}\gamma^* \geq 1/e^t(\tilde{\theta},\hat{\theta}|0) = |\bar{B}_{\theta_0}\bar{P}_{\theta_0}|^{-1}\}\ ,$$

then, from (5.19) and (5.20), we conclude that

$$(5.21)\qquad e(\tilde{\theta},\hat{\theta}|\gamma^*) \geq 1\ ,\quad \forall\ \gamma^* \in S_1\ .$$

Similarly, it follows from (5.4) and (5.8) that

$$(5.22)\qquad e(\underset{\sim}{\theta}^*,\hat{\underset{\sim}{\theta}}|\underset{\sim}{\gamma}^*) = \left\{\frac{|\bar{\underset{\sim}{P}}_{\theta_0} + \underset{\sim}{\gamma}^*\underset{\sim}{\gamma}^{*\prime}|}{|\underset{\sim}{P}_{\theta_0} + a(\Delta^*)(\bar{\underset{\sim}{B}}^{-1}_{\theta_0} - \bar{\underset{\sim}{P}}_{\theta_0}) + b(\Delta^*)\underset{\sim}{\gamma}^*\underset{\sim}{\gamma}^{*\prime}|}\right\}^{1/t},$$

where

$$(5.23)\qquad a(\Delta^*) = 1 - \Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*) \quad \text{and}$$

$$b(\Delta^*) = 2\Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*) - \Pi_{r+4}(\chi^2_{r,\alpha};\Delta^*) .$$

Now $a(\Delta^*) > \alpha$, $\forall\ \Delta^* \geq 0$ and it converges to 1 as $\Delta^* \to \infty$. Also, $0 \leq b(\Delta^*) \leq \Pi_{r+2}(\chi^2_{r,\alpha};\Delta^*) \to 0$ as $\Delta^* \to \infty$. Thus, by (5.12), (5.22) and (5.23), we conclude that for $\underset{\sim}{\gamma}^*$ close to $\underset{\sim}{0}$, (5.22) is less than 1, while it exceeds 1 for all $\underset{\sim}{\gamma}^*$ for which Δ^* is $\geq \Delta^*_0$ where

$$(5.24)\qquad a(\Delta^*_0)[\bar{\underset{\sim}{B}}^{-1}_{\theta_0} - \bar{\underset{\sim}{P}}_{\theta_0}] = [1-b(\Delta^*_0)]\underset{\sim}{\gamma}^*_0\underset{\sim}{\gamma}^{*\prime}_0 \quad \text{and} \quad \underset{\sim}{\gamma}^{*\prime}_0\bar{\underset{\sim}{B}}_{\theta_0}\underset{\sim}{\gamma}^*_0 = \Delta^*_0 .$$

Thus, if we let

$$(5.25)\qquad S_2 = \{\underset{\sim}{\gamma}^* : \underset{\sim}{\gamma}^{*\prime}\bar{\underset{\sim}{B}}_{\theta_0}\underset{\sim}{\gamma}^* \geq \Delta^*_0\} ,$$

then from (5.22)-(5.25), we conclude that

$$(5.26)\qquad e(\underset{\sim}{\theta}^*,\hat{\underset{\sim}{\theta}}|\underset{\sim}{\gamma}^*) \geq 1 \quad \text{for every} \quad \underset{\sim}{\gamma}^* \in S_2 .$$

Finally, we have from (5.2), (5.8) and (5.23),

$$e(\underset{\sim}{\theta}^*,\tilde{\underset{\sim}{\theta}}|\underset{\sim}{\gamma}^*) = \{|\bar{\underset{\sim}{B}}^{-1}_{\theta_0} - \{1-a(\Delta^*)\}\{\bar{\underset{\sim}{B}}^{-1}_{\theta_0} - \bar{\underset{\sim}{P}}_{\theta_0}\} + b(\Delta^*)\underset{\sim}{\gamma}^*\underset{\sim}{\gamma}^{*\prime}||\bar{\underset{\sim}{B}}_{\theta_0}|\}^{-1/t}$$

$$(5.27)\qquad = \{|\underset{\sim}{I} + b(\Delta^*)\bar{\underset{\sim}{B}}_{\theta_0}\underset{\sim}{\gamma}^*\underset{\sim}{\gamma}^{*\prime} - (1-a(\Delta^*))(I - \bar{\underset{\sim}{B}}_{\theta_0}\bar{\underset{\sim}{P}}_{\theta_0})|\}^{-1/t} .$$

Thus, for $\underset{\sim}{\gamma}^*$ close to $\underset{\sim}{0}$, (5.27) exceeds one, while, it is ≤ 1 when

$\underset{\sim}{\gamma}^*$ is away from $\underset{\sim}{0}$ (in the sense that $\Delta^* \to \infty$). Note that both (5.19) and (5.20) go to ∞ as $\Delta^* \to \infty$. But, (5.27) converges to 1 as $\Delta^* \to \infty$. Thus, for large Δ^*, $\{\underset{\sim}{\theta}_n^*\}$ and $\{\underset{\sim}{\tilde{\theta}}_n\}$ have similar performaces (better $\{\underset{\sim}{\hat{\theta}}_n\}$) while for smaller values of Δ^*, (5.16) holds. Combining these with (5.9), (5.16) and (5.17), we conclude that the restricted MLE $\underset{\sim}{\hat{\theta}}_n$, though is asymptotically optimal when H_0: $\underset{\sim}{\theta}_0 \in \omega$, is (asymptotically) biased and its a.d.m. becomes larger (in the sense of trace or generalized variance) when $\underset{\sim}{\gamma}^*$ moves away from $\underset{\sim}{0}$ and this makes it less efficient too. On the other hand, the unrestricted MLE $\underset{\sim}{\hat{\theta}}_n$ remains asymptotically unbiased for $\underset{\sim}{\theta}_0$ irrespective of H_0: $\underset{\sim}{\theta}_0 \in \omega$, but is usually not optimal when H_0 holds. As a compromise, the PTMLE $\underset{\sim}{\theta}_n^*$ performs better than $\underset{\sim}{\tilde{\theta}}_n$ when $\underset{\sim}{\gamma}^*$ is small and has uniformly a lower order of bias than $\underset{\sim}{\hat{\theta}}_n$. For large $\underset{\sim}{\gamma}^*$, it performs better than $\underset{\sim}{\hat{\theta}}_n$ and very similarly to $\underset{\sim}{\tilde{\theta}}_n$. Hence, it can be recommended on the grounds of rebustness against any deviation from H_0: $\underset{\sim}{\theta}_0 \in \omega$. Non-optimality of the preliminary test estimator (under H_0: $\underset{\sim}{\theta}_0 \in \omega$) has been studied by Huntsberger (1955). His conclusions do not remain valid when H_0 does not hold. Indeed, in our setup, we have observed that the PTMLE $\underset{\sim}{\theta}_n^*$ may perform better than either $\underset{\sim}{\hat{\theta}}_n$ or $\underset{\sim}{\tilde{\theta}}_n$, near the boundary of ω, when n is large.

REFERENCES

1. Aitchison, J. and Silvey, S.D. (1958). Maximum likelihood estimation of parameters subject to restraints. *Ann. Math. Statist.* 29, 813-828.

2. Bancroft, T.A. and Han, C.P. (1977). Inference based on conditional specification: a note and a bibliography. *Review Internat. Statist.* 45, 117-128.

3. Feder, P.I. (1968). On the distribution of the log likelihood ratio test statistic when the true parameter is near the boundaries of the hypothesis region. *Ann. Math. Statist.* 39, 2044-2055.

4. Huber, P.J. (1967). The behavior of maximum likelihood estimators under nonstandard conditions. *Proc. 5th Berkeley Symp. Math. Statist. Prob.* 1, 221-234.

5. Huntsberger, D.V. (1955). A generalization of a preliminary test procedure for pooling data. *Ann. Math. Statist.* 26, 734-743.

6. Inagaki, N. (1973). Asymptotic relations between the likelihood estimating function and the maximum likelihood estimator. *Ann. Inst. Statist. Math.* 25, 1-26.

7. Kitagawa, T. (1963). Estimation after preliminary tests of signficance. *Univ. of Calif. Publ. Statist.* 3, 147-186.

8 Silvey, S.D. (1959). The Lagrangian multiplier test. *Ann. Math. Statist.* 30, 389-407.